

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5 (i) APPLICANT: Gorman, Cornelia M.,
Groskreutz, Debyra J.

(ii) TITLE OF INVENTION: Prohormone Convertase Transformed Cells and
Polypeptide Synthesis

10 (iii) NUMBER OF SEQUENCES: 54

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Genentech, Inc.
15 (B) STREET: 460 Point San Bruno Blvd
(C) CITY: South San Francisco
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94080

20 (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
25 (D) SOFTWARE: patin (Genentech)

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE: 01-MAR-1993
30 (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 07/887265
(B) FILING DATE: 22-MAY-1992

35

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: 07/803631
- (B) FILING DATE: 06-DEC-1992

5 (vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: PCT/US92/10621
- (B) FILING DATE: 04-DEC-1992

(viii) ATTORNEY/AGENT INFORMATION:

- 10
- (A) NAME: Adler, Carolyn R.
 - (B) REGISTRATION NUMBER: 32,324
 - (C) REFERENCE/DOCKET NUMBER: 748P3

(ix) TELECOMMUNICATION INFORMATION:

- 15
- (A) TELEPHONE: 415/225-2614
 - (B) TELEFAX: 415/952-9881
 - (C) TELEX: 910/371-7168

(2) INFORMATION FOR SEQ ID NO:1:

20 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2355 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

30 TCTAGATCTA GCTGGTGTGT CTCTGATCTT GCTTCTTTTC TCCCAGCCCT 50

TCCTACTTGT GTGAGAACAA GGTTTGTGAGC CATGGAGCAA AGAGGTTGGA 100

35

CTCTGCAGTG TACTGCTTTC GCCTTCTTTT GCGTTTGGTG TGCACTAAGC 150

AGTGTAAG CAAAGAGGCA GTTGTTAAT GAATGGGCGG CGGAGATCCC 200

5

CGGAGGGCAA GAAGCTGCCT CTGCCATCGC CGAAGAACTG GGGTATGACC 250

TTTTGGGTCA GATTGGATCA CTTGAAAATC ACTATTTATT CAAACACAAA 300

10

AGCCATCCTC GGAGGTCCCG AAGAAGCGCT CTCATATCA CTAAGAGGTT 350

15

ATCTGATGAT GATCGTGTGA CGTGGGCTGA ACAACAGTAT GAAAAAGAGA 400

GAAGTAAACG TTCAGTTCAA AAAGACTCAG CATTGGATCT CTTCAATGAT 450

20

CCAATGTGGA ATCAGCAGTG GTACTTGCAA GATACCAGAA TGACTGCAGC 500

TCTGCCCAAG CTGGACCTTC ATGTAATACC TGTTTGGGAA AAGGGTATTA 550

25

CTGGCAAAGG AGTTGTTATT ACTGTACTGG ATGATGGCTT GGAGTGGAAT 600

30

CACACAGACA TTTATGCCAA TTATGATCCA GAGGCTAGCT ATGATTTTAA 650

CGATAATGAT CATGATCCAT TTCCCCGATA TGATCTCACA AATGAAAACA 700

35

AACATGGAAC AAGATGTGCA GGTGAAATTG CCATGCAAGC AAATAATCAC 750

AAGTGTGGGG TTGGAGTTGC ATATAATTCC AAAGTTGGAG GCATAAGAAT 800

5

GCTGGATGGC ATTGTAAGT ATGCCATTGA GGCTAGTTCA ATTGGATTCA 850

10

ACCCTGGCCA TGTGGATATT TACAGTGCAA GCTGGGGCCC TAATGATGAT 900

GGAAAACTG TGGAGGGGCC TGGCAGACTA GCCCAGAAGG CATTGAATA 950

15

TGGTGTCAA CAGGGGAGAC AAGGGAAAGG CTCCATCTTT GTCTGGGCTT 1000

CAGGGAATGG GGGTCGTCAG GGAGATAACT GTGACTGTGA TGGCTACACA 1050

20

GACAGCATTT ACACCATCTC TATCAGCAGT GCCTCCCAGC AAGGCCTGTC 1100

ACCTTGGTAT GCAGAGAAGT GTTCTTCCAC ATTGGCTACC TCCTACAGCA 1150

25

GTGGTGATTA CACAGACCAG CGAATAACAA GCGCTGACCT GCACAATGAC 1200

30

TGCACAGAGA CCCACACAGG CACCTCGGCT TCAGCACCCC TGGCTGCTGG 1250

TATCTTTGCT CTGGCCTTGG AGGCAAACCC AAATCTTACC TGGAGAGATA 1300

35

TGCAGCATCT GGTGTCTGG ACCTCTGAGT ACGACCCATT GGCCAGTAAC 1350

CCAGGTTGGA AAAAGAATGG GGCAGGCTTG ATGGTGAACA GCCGATTGG 1400

5

ATTTGGCTTG CTAAATGCCA AAGCTCTGGT GGATTGGCT GATCCTCGGA 1450

CCTGGAGAAA TGTGCCTGAG AAGAAAGAAT GTGTTGTAAA AGACAATAAC 1500

10

TTTGAGCCTA GAGCCCTGAA AGCTAATGGA GAAGTAATTG TTGAAATCCC 1550

15

AACAAGAGCT TGTGAAGGAC AAGAAAATGC TATCAAGTCT CTGGAACATG 1600

TGCAATTTGA AGCAACAATT GAATATTCTC GTAGAGGAGA CCTTCATGTC 1650

20

ACACTCACTT CTGCTGTTGG AACCAGCACT GTACTGTTGG CTGAAAGGGA 1700

AAGAGATACA TCCCCCAATG GCTTTAAGAA TTGGGACTTC ATGTCTGTTC 1750

25

ATACATGGGG AGAGAATCCT GTAGGCACCT GGACATTGAA AATTACAGAC 1800

30

ATGTCTGGAA GAATGCAAAA TGAAGGAAGG ATTGTGAACT GGAAGTTGAT 1850

TTTGCATGGG ACATCTTCTC AACCAGAGCA CATGAAGCAG CCCC GTGTGT 1900

35

ACACATCCTA CAATACAGTC CAGAATGACA GGAGAGGAGT GGAAAAGATG 1950

GTGAATGTTG TGGAGAAGCG GCCCACACAA AAGAGCCTGA ATGGCAATCT 2000

5

CCTGGTACCC AAAAACTCCA GCAGCAGCAA TGTGGAGGGT AGAAGGGATG 2050

AGCAGGTACA AGGAACTCCT TCAAAGGCCA TGCTGCGACT CCTACAAAGT 2100

10

GCTTTTAGCA AGAATGCACT TTCAAACAA TCACCAAAGA AGTCTCCAAG 2150

15

TGCAAAGCTC AGCATCCCTT ATGAAAGTTT CTATGAAGCC TTGGAAAAGC 2200

TTAACAAGCC CTCCAAGCTT GAAGGCTCTG AAGACAGTCT GTACAGTGAC 2250

20

TATGTTGATG TATTCTATAA CACAAAACCT TATAAGCATA GAGATGACAG 2300

GCTGCTGCAA GCTCTCATGG ACATCCTAAA TGAGGAGAAT TAAAATAAGG 2350

25

AGCTC 2355

30

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2012 bases

5 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

10

TCTAGATGCA TCTTCCCTCT TCGTCCCCTG CTCCACCACC CTGCGCGCCT 50

15

CACAGCCCCG CTTTTCACCT CCAAAGAAGG ATGGAGGGCG GTTGTGGATC 100

CCAGTGGAAG GCGGCCGGGT TCCTCTTCTG TGTGATGGTT TTTGCGTCTG 150

20

CCGAGAGACC CGTCTTCACG AATCATTTTC TTGTGGAGTT GCATAAAGAC 200

GGAGAGGAAG AGGCTCGCCA AGTTGCAGCA GAACACGGCT TTGGAGTCCG 250

25

AAAGCTCCCC TTTGCAGAAG GCCTGTATCA CTTTATCAC AATGGGCTTG 300

30

CAAAGGCCAA AAGAAGACGC AGCCTACACC ATAAGCGGCA GCTAGAGAGA 350

GACCCCAGGA TAAAGATGGC GCTGCAACAA GAAGGATTTG ACCGTAAAAA 400

35

GAGAGGGTAC AGGGACATCA ATGAGATTGA CATCAACATG AATGATCCTC 450

TCTTTACAAA GCAATGGTAC CTGTTCAACA CTGGGCAAGC CGATGGAACT 500

5

CCTGGGCTAG ACTTGAACGT GGCCGAAGCC TGGGAGCTGG GATACACAGG 550

AAAAGGAGTG ACCATTGGAA TTATGGATGA TGAATTGAC TATCTCCACC 600

10

CAGACCTGGC CTACAACTAC AACGCTGATG CAAGTTATGA CTCAGCAGC 650

15

AATGACCCCT ACCCATACCC TCGATACACA GATGACTGGT TCAACAGCCA 700

TGGAAGTAGG TGTGCAGGAG AAGTTTCTGC TGCAGCCAGC AACAATATCT 750

20

GTGGAGTCGG CGTAGCATAC AACTCCAAGG TGGCAGGGAT CCGGATGCTG 800

GACCAGCCCT TTATGACAGA CATCATCGAA GCCTCCTCCA TCAGCCACAT 850

25

GCCTCAACTG ATCGACATCT ACAGTGCAAG CTGGGGCCCC ACAGACAATG 900

30

GGAAGACGGT TGATGGGCCC CGAGAGCTCA CACTCCAGGC CATGGCTGAT 950

GGCGTGAACA AGGGCCGTGG GGGCAAAGGC AGCATCTATG TGTGGGCCTC 1000

35

TGGGGACGGT GGCAGCTACG ATGACTGCAA CTGTGACGGC TATGCTTCAA 1050

GCATGTGGAC CATCTCCATC AACTCAGCCA TCAATGATGG CAGGACTGCC 1100

5

TTGTATGATG AGAGTTGCTC TTCCACCTTA GCATCCACCT TCAGCAATGG 1150

GAGGAAGAGG AATCCTGAGG CTGGTGTGGC TACCACAGAC TTGTATGGCA 1200

10

ACTGTACTCT GAGACACTCT GGGACATCTG CAGCTGCTCC GGAGGCAGCT 1250

15

GGCGTGTTTG CATTAGCTTT GGAGGCTAAC CTGGATCTGA CCTGGAGAGA 1300

CATGCAACAT CTGACTGTGC TCACCTCCAA GCGGAACCAG CTTTCATGATG 1350

20

AGGTTTCATCA GTGGCGACGG AATGGGGTTG GCCTGGAATT TAATCACCTC 1400

TTTGGCTACG GAGTCCTTGA TGCAGGTGCC ATGGTGAAAA TGGCTAAAGA 1450

25

CTGGAAAAC TGTCCGGAGA GATTCCATTG TGTGGGAGGC TCTGTGCAGA 1500

30

ACCCTGAAAA AATACCACCC ACCGGCAAGC TGGTACTGAC CCTCAAAACA 1550

AATGCATGTG AGGGGAAGGA AAAC TTCGTC CGCTACCTGG AGCACGTCCA 1600

35

AGCTGTCATC ACAGTCAACG CGACCAGGAG AGGAGACCTG AACATCAACA 1650

TGACCTCCCC AATGGGCACC AAGTCCATTT TGCTGAGCCG GCGTCCCAGA 1700

5

GACGACGACT CCAAGGTGGG CTTTGACAAG TGGCCTTTCA TGACCACCCA 1750

CACCTGGGGG GAGGATGCCC GAGGGACCTG GACCCTGGAG CTGGGGTTTG 1800

10

TGGGCAGTGC ACCACAGAAG GGGTTGCTGA AGGAATGGAC CCTGATGCTT 1850

15

CACGGCACAC AGAGCGCCCC ATACATCGAT CAGGTGGTGA GGGATTACCA 1900

GTCTAAGCTG GCCATGTCCA AGAAGCAGGA GCTGGAGGAA GAGCTGGATG 1950

20

AGGCTGTGGA GAGAAGCCTG CAAAGTATCC TGAGAAAGAA CTAGGGCCAC 2000

GCTTCCGAAT TC 2012

25

(2) INFORMATION FOR SEQ ID NO:3:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 753 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Glu Gln Arg Gly Trp Thr Leu Gln Cys Thr Ala Phe Ala Phe
1 5 10 15
5
Phe Cys Val Trp Cys Ala Leu Ser Ser Val Lys Ala Lys Arg Gln
20 25 30
Phe Val Asn Glu Trp Ala Ala Glu Ile Pro Gly Gly Gln Glu Ala
10 35 40 45
Ala Ser Ala Ile Ala Glu Glu Leu Gly Tyr Asp Leu Leu Gly Gln
50 55 60
15 Ile Gly Ser Leu Glu Asn His Tyr Leu Phe Lys His Lys Ser His
65 70 75
Pro Arg Arg Ser Arg Arg Ser Ala Leu His Ile Thr Lys Arg Leu
80 85 90
20 Ser Asp Asp Asp Arg Val Thr Trp Ala Glu Gln Gln Tyr Glu Lys
95 100 105
Glu Arg Ser Lys Arg Ser Val Gln Lys Asp Ser Ala Leu Asp Leu
25 110 115 120
Phe Asn Asp Pro Met Trp Asn Gln Gln Trp Tyr Leu Gln Asp Thr
125 130 135
30 Arg Met Thr Ala Ala Leu Pro Lys Leu Asp Leu His Val Ile Pro
140 145 150
Val Trp Glu Lys Gly Ile Thr Gly Lys Gly Val Val Ile Thr Val
155 160 165
35

	Leu Asp Asp Gly Leu Glu Trp Asn His Thr Asp Ile Tyr Ala Asn	170	175	180
5	Tyr Asp Pro Glu Ala Ser Tyr Asp Phe Asn Asp Asn Asp His Asp	185	190	195
	Pro Phe Pro Arg Tyr Asp Leu Thr Asn Glu Asn Lys His Gly Thr	200	205	210
10	Arg Cys Ala Gly Glu Ile Ala Met Gln Ala Asn Asn His Lys Cys	215	220	225
	Gly Val Gly Val Ala Tyr Asn Ser Lys Val Gly Gly Ile Arg Met	230	235	240
15	Leu Asp Gly Ile Val Thr Asp Ala Ile Glu Ala Ser Ser Ile Gly	245	250	255
	Phe Asn Pro Gly His Val Asp Ile Tyr Ser Ala Ser Trp Gly Pro	260	265	270
20	Asn Asp Asp Gly Lys Thr Val Glu Gly Pro Gly Arg Leu Ala Gln	275	280	285
	Lys Ala Phe Glu Tyr Gly Val Lys Gln Gly Arg Gln Gly Lys Gly	290	295	300
	Ser Ile Phe Val Trp Ala Ser Gly Asn Gly Gly Arg Gln Gly Asp	305	310	315
30	Asn Cys Asp Cys Asp Gly Tyr Thr Asp Ser Ile Tyr Thr Ile Ser	320	325	330
	Ile Ser Ser Ala Ser Gln Gln Gly Leu Ser Pro Trp Tyr Ala Glu	335	340	345
35				

	Lys Cys Ser Ser Thr Leu Ala Thr Ser Tyr Ser Ser Gly Asp Tyr	350	355	360
5	Thr Asp Gln Arg Ile Thr Ser Ala Asp Leu His Asn Asp Cys Thr	365	370	375
	Glu Thr His Thr Gly Thr Ser Ala Ser Ala Pro Leu Ala Ala Gly	380	385	390
10	Ile Phe Ala Leu Ala Leu Glu Ala Asn Pro Asn Leu Thr Trp Arg	395	400	405
	Asp Met Gln His Leu Val Val Trp Thr Ser Glu Tyr Asp Pro Leu	410	415	420
15	Ala Ser Asn Pro Gly Trp Lys Lys Asn Gly Ala Gly Leu Met Val	425	430	435
	Asn Ser Arg Phe Gly Phe Gly Leu Leu Asn Ala Lys Ala Leu Val	440	445	450
20	Asp Leu Ala Asp Pro Arg Thr Trp Arg Asn Val Pro Glu Lys Lys	455	460	465
	Glu Cys Val Val Lys Asp Asn Asn Phe Glu Pro Arg Ala Leu Lys	470	475	480
	Ala Asn Gly Glu Val Ile Val Glu Ile Pro Thr Arg Ala Cys Glu	485	490	495
30	Gly Gln Glu Asn Ala Ile Lys Ser Leu Glu His Val Gln Phe Glu	500	505	510
	Ala Thr Ile Glu Tyr Ser Arg Arg Gly Asp Leu His Val Thr Leu	515	520	525
35				

	Thr Ser Ala Val Gly Thr Ser Thr Val Leu Leu Ala Glu Arg Glu		
	530	535	540
5	Arg Asp Thr Ser Pro Asn Gly Phe Lys Asn Trp Asp Phe Met Ser	545	555
	Val His Thr Trp Gly Glu Asn Pro Val Gly Thr Trp Thr Leu Lys	560	570
10	Ile Thr Asp Met Ser Gly Arg Met Gln Asn Glu Gly Arg Ile Val	575	585
	Asn Trp Lys Leu Ile Leu His Gly Thr Ser Ser Gln Pro Glu His	590	600
15	Met Lys Gln Pro Arg Val Tyr Thr Ser Tyr Asn Thr Val Gln Asn	605	615
	Asp Arg Arg Gly Val Glu Lys Met Val Asn Val Val Glu Lys Arg	620	630
20	Pro Thr Gln Lys Ser Leu Asn Gly Asn Leu Leu Val Pro Lys Asn	635	645
	Ser Ser Ser Ser Asn Val Glu Gly Arg Arg Asp Glu Gln Val Gln	650	660
	Gly Thr Pro Ser Lys Ala Met Leu Arg Leu Leu Gln Ser Ala Phe	665	675
30	Ser Lys Asn Ala Leu Ser Lys Gln Ser Pro Lys Lys Ser Pro Ser	680	690
	Ala Lys Leu Ser Ile Pro Tyr Glu Ser Phe Tyr Glu Ala Leu Glu	695	705
35		700	

Lys Leu Asn Lys Pro Ser Lys Leu Glu Gly Ser Glu Asp Ser Leu
710 715 720

5 Tyr Ser Asp Tyr Val Asp Val Phe Tyr Asn Thr Lys Pro Tyr Lys
725 730 735

His Arg Asp Asp Arg Leu Leu Gln Ala Leu Met Asp Ile Leu Asn
740 745 750

10 Glu Glu Asn
753

(2) INFORMATION FOR SEQ ID NO:4:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 637 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Glu Gly Gly Cys Gly Ser Gln Trp Lys Ala Ala Gly Phe Leu
1 5 10 15

25 Phe Cys Val Met Val Phe Ala Ser Ala Glu Arg Pro Val Phe Thr
20 25 30

Asn His Phe Leu Val Glu Leu His Lys Asp Gly Glu Glu Glu Ala
35 40 45

30 Arg Gln Val Ala Ala Glu His Gly Phe Gly Val Arg Lys Leu Pro
50 55 60

Phe Ala Glu Gly Leu Tyr His Phe Tyr His Asn Gly Leu Ala Lys
35 65 70 75

	Ala Lys Arg Arg Arg Ser Leu His His Lys Arg Gln Leu Glu Arg		
	80	85	90
5	Asp Pro Arg Ile Lys Met Ala Leu Gln Gln Glu Gly Phe Asp Arg	95	100 105
	Lys Lys Arg Gly Tyr Arg Asp Ile Asn Glu Ile Asp Ile Asn Met	110	115 120
10	Asn Asp Pro Leu Phe Thr Lys Gln Trp Tyr Leu Phe Asn Thr Gly	125	130 135
	Gln Ala Asp Gly Thr Pro Gly Leu Asp Leu Asn Val Ala Glu Ala	140	145 150
15	Trp Glu Leu Gly Tyr Thr Gly Lys Gly Val Thr Ile Gly Ile Met	155	160 165
	Asp Asp Gly Ile Asp Tyr Leu His Pro Asp Leu Ala Tyr Asn Tyr	170	175 180
20	Asn Ala Asp Ala Ser Tyr Asp Phe Ser Ser Asn Asp Pro Tyr Pro	185	190 195
	Tyr Pro Arg Tyr Thr Asp Asp Trp Phe Asn Ser His Gly Thr Arg	200	205 210
	Cys Ala Gly Glu Val Ser Ala Ala Ala Ser Asn Asn Ile Cys Gly	215	220 225
30	Val Gly Val Ala Tyr Asn Ser Lys Val Ala Gly Ile Arg Met Leu	230	235 240
	Asp Gln Pro Phe Met Thr Asp Ile Ile Glu Ala Ser Ser Ile Ser	245	250 255
35			

	His Met Pro Gln Leu Ile Asp Ile Tyr Ser Ala Ser Trp Gly Pro	260	265	270
5	Thr Asp Asn Gly Lys Thr Val Asp Gly Pro Arg Glu Leu Thr Leu	275	280	285
	Gln Ala Met Ala Asp Gly Val Asn Lys Gly Arg Gly Gly Lys Gly	290	295	300
10	Ser Ile Tyr Val Trp Ala Ser Gly Asp Gly Gly Ser Tyr Asp Asp	305	310	315
	Cys Asn Cys Asp Gly Tyr Ala Ser Ser Met Trp Thr Ile Ser Ile	320	325	330
15	Asn Ser Ala Ile Asn Asp Gly Arg Thr Ala Leu Tyr Asp Glu Ser	335	340	345
	Cys Ser Ser Thr Leu Ala Ser Thr Phe Ser Asn Gly Arg Lys Arg	350	355	360
20	Asn Pro Glu Ala Gly Val Ala Thr Thr Asp Leu Tyr Gly Asn Cys	365	370	375
25	Thr Leu Arg His Ser Gly Thr Ser Ala Ala Ala Pro Glu Ala Ala	380	385	390
	Gly Val Phe Ala Leu Ala Leu Glu Ala Asn Leu Asp Leu Thr Trp	395	400	405
30	Arg Asp Met Gln His Leu Thr Val Leu Thr Ser Lys Arg Asn Gln	410	415	420
	Leu His Asp Glu Val His Gln Trp Arg Arg Asn Gly Val Gly Leu	425	430	435
35				

	Glu Phe Asn His Leu Phe Gly Tyr Gly Val Leu Asp Ala Gly Ala		
	440	445	450
5	Met Val Lys Met Ala Lys Asp Trp Lys Thr Val Pro Glu Arg Phe	455	460 465
	His Cys Val Gly Gly Ser Val Gln Asn Pro Glu Lys Ile Pro Pro	470	475 480
10	Thr Gly Lys Leu Val Leu Thr Leu Lys Thr Asn Ala Cys Glu Gly	485	490 495
	Lys Glu Asn Phe Val Arg Tyr Leu Glu His Val Gln Ala Val Ile	500	505 510
15	Thr Val Asn Ala Thr Arg Arg Gly Asp Leu Asn Ile Asn Met Thr	515	520 525
	Ser Pro Met Gly Thr Lys Ser Ile Leu Leu Ser Arg Arg Pro Arg	530	535 540
20	Asp Asp Asp Ser Lys Val Gly Phe Asp Lys Trp Pro Phe Met Thr	545	550 555
	Thr His Thr Trp Gly Glu Asp Ala Arg Gly Thr Trp Thr Leu Glu	560	565 570
	Leu Gly Phe Val Gly Ser Ala Pro Gln Lys Gly Leu Leu Lys Glu	575	580 585
30	Trp Thr Leu Met Leu His Gly Thr Gln Ser Ala Pro Tyr Ile Asp	590	595 600
	Gln Val Val Arg Asp Tyr Gln Ser Lys Leu Ala Met Ser Lys Lys	605	610 615
35			

Gln Glu Leu Glu Glu Glu Leu Asp Glu Ala Val Glu Arg Ser Leu
620 625 630

5 Gln Ser Ile Leu Arg Lys Asn
635 637

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 32 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCAAAATCTA GAYKGCNATY GTNGAYGAKG GN 32

20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 30 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

30

AAGCATGAGC TCNGGRGCRG CRGCNGANCC 30

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GATATCACTC AGATCGATGA ATTCGAGCTC 30

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AAGCTTTCTA GAGGATCCCT CTGGTGGATT TGG 33

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AAGCTTGAAT TCTCCAACCC CACACTTGTG 30

5

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGATCGATGA ATTCGAGCTC 20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CATTCTCGAA AAAAGAGACA A 21

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CATTCTGTAA AAAAGAGACA A 21

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CATTCTAGAG CAAAGAGACA A 21

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CATTCTAGAA AAGCAAGACA A 21

5

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 21 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CATTCTAGAA AAAGAGCACA A 21

20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

25

(A) LENGTH: 21 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ACCTGGAGCA AAGCTTCTCT G 21

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 bases

5 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

10

ACCTGGAGCG CTAGGTCTCT G 21

15

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CATAAGCTTA CCATGGCCCT GTGGATGCGC 30

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CATTCTAGAC TAGTTGCAGT AGTTCTCCAG 30

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Lys Thr Arg Arg

1 4

5

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Leu Gln Lys Arg

1 4

15

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Arg Thr Lys Arg

1 4

25

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Arg Gln Lys Arg

1 4

5

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Lys Thr Lys Arg

1 4

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CTCTGCCTCC CGCTTGGTCC TGGGTGTGTA G 31

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

5

CACGCTTCTG CCGGGATCCC TC 22

10 (2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 bases

(B) TYPE: nucleic acid

15 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

20

CTCTGCCTCC CGCTTGGTCT TCGGTGTGTA G 31

(2) INFORMATION FOR SEQ ID NO:28:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

35 GATATGAAGA GCAGATCTTT TGGACCTCCG AGGATG 36

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 39 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CTTATGGTGT AAGCTTCGTT TTGCTCTGGC CTTTGCAAG 39

15 (2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 21 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

25

TACAACTCAC CGCGGGTCCT G 21

30 (2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 38 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

5 AAGATGGGAT GGGATGATGA CCGTTTCCGC CTTGATGT 38

(2) INFORMATION FOR SEQ ID NO:32:

10 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ACATCACGGC GGAAACGGTC ATCATCCCAT CCCATCTT 38

20

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

25

(A) LENGTH: 28 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GATATAAGCT TGAGAGTGTA GAAGGGGC 28

35

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGCTTGACAT CATTGGCTGA CACTTTCGAA CACATGATAG 40

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AAGATGGGAT GGGATGATGA GCGCCGGACC CTCATGGACA T 41

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

ATGTCCATGA GGGTCCGGCG CTCATCATCC CATCCCATCT T 41

5

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 28 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GATATAAGCT TGAGAGTGTA GAAGGGGC 28

20

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

25

(A) LENGTH: 51 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TGAGCGACAG CACCCCTTG GAGCCCCCGC CCTTGATCT CATGGAGGAT 50

35 T 51

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 51 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

TCAGCTCCCC TCGTCGGGCG GGGTCCGAGT GCCGTTTCCG CCGTGATGTT 50

15

C 51

(2) INFORMATION FOR SEQ ID NO:40:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

30 ACGTGGGCAG CCCCGTGGTG GCGAACAGAA CATCACGGCG GAAACGGC 48

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

TGTTGCCAC CACGGGGCTG CCCACGTAAT CCTCCATGAG ATACAAGG 48

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CTCGGACCCC GCCCGACGAG GGGAGC 26

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

5 GCGGGGGCTC CAAGGGGGTG CTGTCGC 27

(2) INFORMATION FOR SEQ ID NO:44:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

15

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

20 GTAGAACAAC ATGGACATGG TGC CAATATT GTCGACTCTG GAGTCGACCT 50

GCAG 54

25

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 bases

30

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

35

CACATAAAAC AAGATGGACA TGGTCTTGTT CACCTGTAGG ATCCCCGG 48

(2) INFORMATION FOR SEQ ID NO:46:

5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

15 AGTAAGGAAA AGGATGGTCA TGGTGGAGGT CGACAAGCTT GAGAATTCAA 50

TCG 53

20

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

30 Xaa Xaa Xaa Arg

1

4

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

5 (B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

10 His Ser Arg Lys Lys Arg Gln
1 5 7

(2) INFORMATION FOR SEQ ID NO:49:

15 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

His Ser Val Lys Lys Arg Gln
1 5 7

25 (2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

35 His Ser Arg Ala Lys Arg Gln
1 5 7

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

5 (B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

10 His Ser Arg Lys Ala Arg Gln

1 5 7

(2) INFORMATION FOR SEQ ID NO:52:

15 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

His Ser Arg Lys Arg Ala Gln

1 5 7

25 (2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Thr Trp Ser Lys Ala Ser Gln

35 1 5 7

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

5 (B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

10 Thr Trp Ser Ala Arg Ser Gln

1 5 7